Metabarcoding and Metagenomics 2: 1–15
DOI 10.3897/mbmg.2.23297

Research Article

Quantitative monitoring of multispecies fish environmental DNA using high-throughput sequencing

Masayuki Ushio[1,2,3,4], Hiroaki Murakami[5], Reiji Masuda[5], Tetsuya Sado[6], Masaki Miya[6], Sho Sakurai[7], Hiroki Yamanaka[3], Toshifumi Minamoto[8], Michio Kondoh[3]

1 *Center for Ecological Research, Kyoto University, Otsu, Japan.*
2 *PRESTO, Japan Science and Technology Agency, Kawaguchi, Japan.*
3 *Department of Environmental Solution Technology, Faculty of Science and Technology, Ryukoku University, Otsu, Japan.*
4 *Joint Research Center for Science and Technology, Ryukoku University, Otsu, Japan.*
5 *Maizuru Fisheries Research Station, Kyoto University, Maizuru, Japan.*
6 *Department of Ecology and Environmental Sciences, Natural History Museum and Institute Chiba, Chiba, Japan.*
7 *Graduate School of Science and Technology, Ryukoku University, Otsu, Japan.*
8 *Graduate School of Human Development and Environment, Kobe University, Kobe, Japan.*

Corresponding author: *Masayuki Ushio* (ong8181@gmail.com)

Academic editor: *Andrew R. Mahon* | **Received** 28 December 2017 | **Accepted** 21 February 2018 | **Published** 14 March 2018

Abstract

Effective ecosystem conservation and resource management require quantitative monitoring of biodiversity, including accurate descriptions of species composition and temporal variations of species abundance. Accordingly, quantitative monitoring of biodiversity has been performed for many ecosystems, but it is often time- and effort- consuming and costly. Recent studies have shown that environmental DNA (eDNA), which is released to the environment from macro-organisms living in a habitat, contains information about species identity and abundance. Thus, analysing eDNA would be a promising approach for more efficient biodiversity monitoring. In the present study, internal standard DNAs (i.e. known amounts of short DNA fragments from fish species that have never been observed in a sampling area) were added to eDNA samples, which were collected weekly from a coastal marine ecosystem in Maizuru Bay, Japan (from April 2015 to March 2016) and metabarcoding analysis was performed using Illumina MiSeq to simultaneously identify fish species and quantify fish eDNA copy numbers. A correction equation was obtained for each sample using the relationship between the number of sequence reads and the added amount of the standard DNAs and this equation was used to estimate the copy numbers from the sequence reads of non-standard fish eDNA. The calculated copy numbers showed significant positive correlation with those determined by quantitative PCR, suggesting that eDNA metabarcoding with standard DNA enabled useful quantifications of eDNA. Furthermore, for samples that show a high level of PCR inhibition, this method might allow more accurate quantification than qPCR because the correction equations generated using internal standard DNAs would include the effect of PCR inhibition. A single run of Illumina MiSeq produced >70 quantitative fish eDNA time series in this study, showing that this method could contribute to more efficient and quantitative monitoring of biodiversity.

Keywords

biodiversity monitoring; environmental DNA; DNA quantification; internal standard DNA; qPCR; metabarcoding

Introduction

Effective ecosystem conservation and resource management require quantitative monitoring of biodiversity, including accurate descriptions of species composition and temporal variations of species abundance. Accordingly, quantitative monitoring of biodiversity has often been performed for many ecosystems. For example, fishing (in aquatic ecosystems), the camera/video trap method (in terrestrial ecosystems) and direct visual census (in aquat-

Copyright *Masayuki Ushio et al.* This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ic and terrestrial ecosystems) have traditionally been used as tools for biodiversity monitoring (see, for example, Masuda et al. 2010, Samejima et al. 2012). These data are invaluable in conservation ecology, but at the same time, the traditional approaches are usually time- and effort-consuming and costly. In addition, most of the traditional methods require professional expertise such as taxonomic identification skill in the field. These difficulties prevent the collection of quantitative, comprehensive (i.e. multispecies and fine-time resolution) and long-term monitoring data about biodiversity.

Environmental DNA (eDNA), which designates DNA isolated from environmental samples (e.g. water or soil) without sampling target (macro-)organism(s), has been used to detect the presence of macro-organisms, particularly those living in an aquatic environment (e.g. Taberlet et al. 2012, Miya et al. 2015, Yamamoto et al. 2017). In the case of macro-organisms, eDNA originates from various sources such as metabolic waste, damaged tissue, dead individuals and/or spawning events (Kelly et al. 2014, Barnes and Turner 2016) and the eDNA contains information about the species identity of organisms that produced it. Since the first application of eDNA analysis to natural ecosystems (Ficetola et al. 2008), eDNA in aquatic ecosystems has been used in many studies as a tool for investigation of the distributions of fish species in ponds, rivers and seawater (Jerde et al. 2011, Minamoto et al. 2012, Takahara et al. 2012, 2013, Sigsgaard et al. 2015, Simmons et al. 2016), as well as the distributions of other aquatic/semiaquatic/terrestrial organisms (Baschien et al. 2008, Davy et al. 2015, Fukumoto et al. 2015, Deiner et al. 2016, Bista et al. 2017, Ushio et al. 2017a, b). Recently, researchers have begun to apply high-throughput sequencing technology (e.g. Illumina MiSeq) and universal primer sets to eDNA studies (Taberlet et al. 2012, Kelly et al. 2014, Miya et al. 2015, Simmons et al. 2016, Ushio et al. 2017a, b, Yamamoto et al. 2017). A previous study demonstrated that an eDNA metabarcoding approach using fish-targeting universal primers (MiFish primers) enabled detection of more than 230 fish species from seawater in a single study (Miya et al. 2015). Accordingly, the eDNA metabarcoding approach has become a cost- and labour-effective approach for estimating aquatic biodiversity.

Though the eDNA metabarcoding approach has greatly improved the efficiency of biodiversity monitoring, several potential limitations prevent its use as a tool for quantitative monitoring of biodiversity. First, whether the quantity of eDNA is a reliable index of the abundance (or biomass) of macro-organisms is still challenging despite many reports of positive correlations between them (Takahara et al. 2012, Yamamoto et al. 2016, Barnes and Turner 2016, Klobucar et al. 2017, Stoeckle et al. 2017). Second, even if the quantity of eDNA is an index of the abundance/biomass of macro-organisms, the number of eDNA sequence reads obtained by high-throughput sequencing may not be an index of the quantity of eDNA partly due to several experimental problems such as PCR inhibitions (i.e. the concentrations of PCR inhibitors depend on samples; Schrader et al. 2012) and thus the quantity of eDNA in an environment cannot be estimated by the eDNA metabarcoding approach.

Regarding the first issue, some studies showed that the eDNA quantity could be a proxy for the abundance or biomass of macro-organisms under particular conditions such as tank and mesocosm experiments (e.g. Thomsen et al. 2011, Takahara et al. 2012). In addition, a recent study showed that eDNA quantity is a proxy for the abundance of fish even in an open ocean ecosystem if appropriate spatial information is incorporated (Yamamoto et al. 2016). Thus, there have been many reports of positive and significant linear relationships between eDNA quantity and the abundance/biomass of macro-organisms (Thomsen et al. 2011, Takahara et al. 2012, Evans et al. 2015, Thomas et al. 2015, Yamamoto et al. 2016, Klobucar et al. 2017, Stoeckle et al. 2017). Nonetheless, the general use of eDNA as a proxy of fish abundance/biomass is still controversial because factors that influence eDNA quantity, such as eDNA decay rates in an environment and their release rates from target organisms, are likely to depend on the ecology and physiology of target species and other biotic/abiotic factors (Dejean et al. 2011, Maruyama et al. 2014, Barnes and Turner 2016, Tsuji et al. 2017) and because the eDNA quantity found in a sample could also be influenced by water flow in an environment (i.e. eDNA transport). The findings of such studies imply that an accurate estimation of organism abundance/biomass requires sample-specific calibrations that appropriately take into account biotic/abiotic factors (e.g. fish physiological conditions, water temperature, water flow and spatial information). Altogether, the above evidence suggested that information about the abundance/biomass is "encoded" in the quantity of eDNA at least to some extent and that the quantity of eDNA may be used as "a rough index" for abundance/biomass, but careful interpretations are necessary especially when other related information (e.g. physicochemical properties of water and the ecology of target species) is not available.

Regarding the second issue, some potential approaches to solve this problem have been reported in the field of microbial ecology. For example, Smets et al. (2016) added an internal standard DNA (DNA of a microbial species that had never been found in a sample) of known quantity to a soil sample. They used the number of sequence reads of the internal standard DNA to estimate the sequence reads per number of DNA copies and converted the sequence reads of DNAs from unknown microbial species (i.e. non-standard microbial species) to the number of microbial DNA copies. The total number of microbial DNA copies estimated was significantly positively correlated with other reliable and quantitative indices of soil microbial abundance. Although the quality and quantity of microbial DNA from soil samples could be different from those of macrobial (e.g.

fish) eDNA from water samples and, although laboratory based biases such as DNA extraction methods and PCR amplification biases could influence the results, this approach is potentially useful to resolve the second issue.

In the present study, we focused on the second issue (i.e. the quantification of eDNA using high-throughput sequencing) and did not explicitly try to resolve the first issue (i.e. the accurate estimation of species abundance/biomass based on eDNA quantity), because at present various biotic/abiotic factors at fine spatiotemporal resolutions, for some of which data are not currently available in the study region, should be incorporated to resolve this first issue (e.g. Yamamoto et al. 2016). The internal standard DNA method was applied to the eDNA metabarcoding approach to enable quantitative monitoring of multispecies fish eDNA in a coastal marine ecosystem (i.e. identification of fish species and quantification of the number of fish eDNA copies simultaneously). Water samples were collected weekly from a sampling station in Maizuru Bay, located on the Japan Sea coast of central Japan and eDNAs were extracted from the samples. Known quantities of short DNA fragments derived from five fish species that have never been observed in the sampling region (freshwater fish species in Southeast Asia or Africa) were added as internal standard DNAs to each eDNA sample. Using the relationships between the quantity and sequence reads (generated by Illumina MiSeq) of the internal standard DNAs, the sequence reads were converted to the calculated DNA copy numbers. The reliability of this internal standard DNA method was tested by comparing the calculated DNA copy numbers with those quantified by quantitative PCR (qPCR) using several statistical methods. Specifically, the followings were tested: 1) whether numbers of sequence reads of the internal standard DNAs linearly correlate with their quantity (copy numbers), 2) whether there is a positive and significant relationship between the calculated DNA copy numbers and those quantified by qPCR and 3) whether temporal dynamics shown by the internal standard method are comparable with those shown by qPCR.

Material and methods

Study site

Water samples were collected at a floating pier in the Maizuru Fishery Research Station of Kyoto University (Nagahama, Maizuru, Kyoto, Japan: 35°29′24.66” N, 135°22′5.76” E; Fig. 1). The sampling point was located 11 m from the shore, with a bottom depth of 4 m. The adjacent area included a rocky reef, brown algae macrophyte and filamentous epiphyte vegetation, live oysters, *Crassostrea gigas* (Thunberg, 1793) and their shells, a sandy or muddy silt bottom and an artificial vertical structure that functioned as a fish reef. The surface water temperature and salinity in the area ranged from 1.2 to 30.8°C and from 4.14 to 34.09 ‰, respectively. The mean (±SD) surface salinity was 30.0 ± 2.9 ‰ (n = 1,753) and did not show clear seasonality. Further information on the study area is available in Masuda (2008) and Masuda et al. (2010).

Water sampling and DNA extraction

All sampling and filtering equipment was washed with a 10% commercial bleach solution before use. One thousand millilitres of seawater were collected once a week from 7 April 2015 to 29 March 2016 from a pier (Fig. 1b) in the study area using a polyethylene bottle. Thus, the total number of eDNA samples (excluding artificial seawater samples as negative controls) was 52. The collected water samples were immediately taken to the laboratory and filtered using 47-mm diameter glass-fibre filters (nominal pore size, 0.7 µm; Whatman, Maidstone, UK). The sampling bottles were gently shaken before the filtration. After the filtration, each filter was wrapped in commercially available aluminium foil and stored at – 20°C before eDNA extraction. Artificial seawater (1,000 ml), prepared from distilled water (Water Purifiers WG202, Yamato Scientific, Tokyo, Japan), was used as the field negative control and sampling bottles filled with artificial seawater were treated identically to the eDNA samples in order to monitor contamination during the bottle handling, water filtering and subsequent DNA extraction. The field negative controls were obtained at the first sampling event in each month (a total of 12 field negative controls during the one-year sampling period) and all negative controls produced a negligible number of sequences (i.e. after removing reads of standard DNAs, the average number of sequence reads to which species names were assigned was 2,305 for environmental samples, while it was 21 for negative controls; see Suppl. material 1).

DNA was extracted from the filters using a DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany) in combination with a spin column (EZ-10; Bio Basic, Markham, Ontario, Canada). After removal of the attached membrane from the spin column (EZ-10), the filter was tightly folded into a small cylindrical shape and placed in the spin column. The spin column was centrifuged at 6,000 g for 1 min to remove excess water from the filter. The column was then placed in the same 2-ml tube and subjected to cell lysis using proteinase K. For the lysis, sterilised H_2O (200 µl), proteinase K (10 µl) and buffer AL (100 µl) were mixed and the mixed solution was gently pipetted on to the folded filter in the spin column. The column was then placed on a 56°C preheated aluminium heat block and incubated for 30 min. After the incubation, the spin column was centrifuged at 6,000 g for 1 min to collect DNA. In order to increase the yield of DNA from the filter, 200 µl of sterilised TE buffer was gently pipetted on to the folded filter and the spin column was again centrifuged at 6,000 g for 1 min. The collected DNA solution (about 500 µl) was purified using a DNeasy Blood and

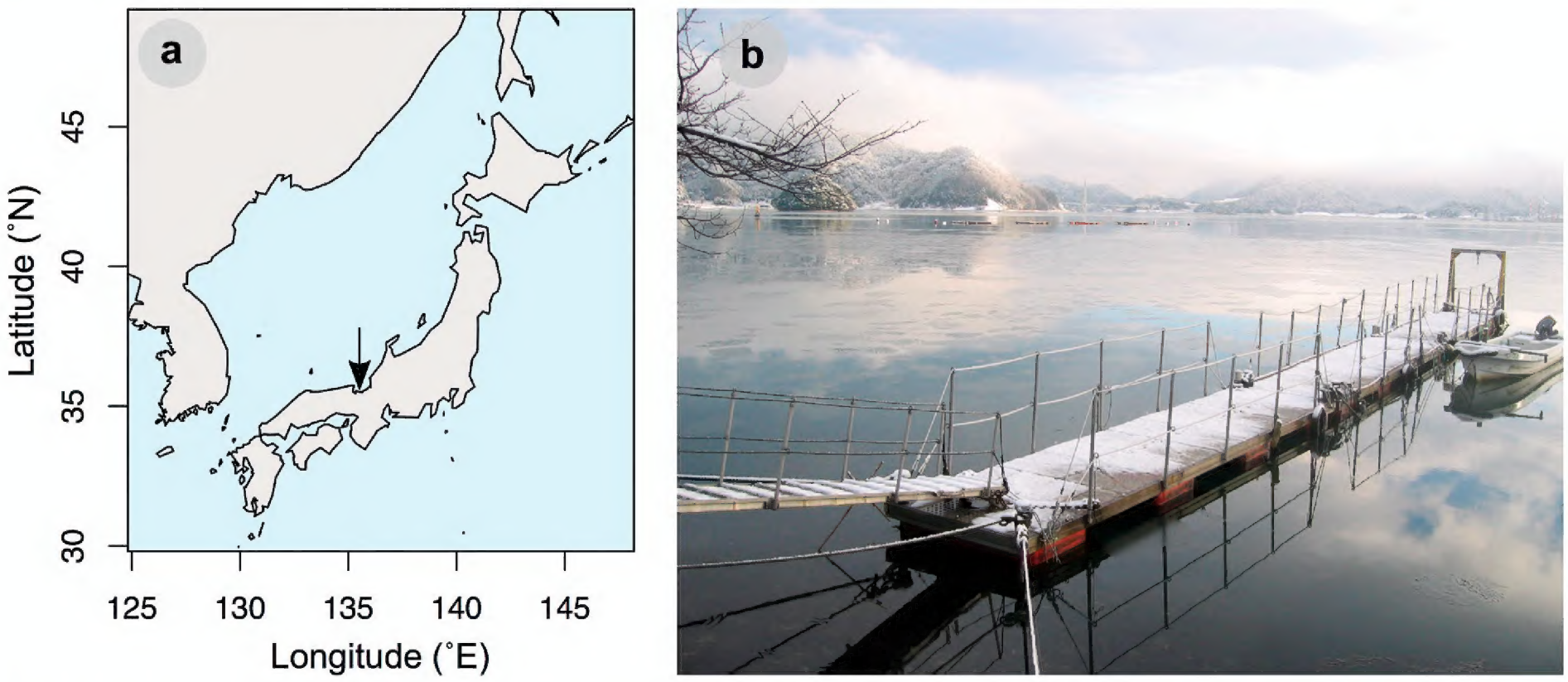

Figure 1. Location of the research site (**a**). The arrow indicates the research site. A floating pier in the Maizuru Fishery Research Station of Kyoto University, Maizuru, Kyoto, Japan, where the weekly water sampling was performed. (**b**). Photo taken in winter by R. Masuda.

Tissue Kit following the manufacturer's protocol. After the purification steps, DNA was eluted with the elution buffer (100 µl) provided in the kit.

Preparation of standard fish DNAs

Extracted DNAs of five fish species, *Saurogobio immaculatus* Koller, 1927, *Elopichthys bambusa* (Richardson, 1845), *Carassioides acuminatus* (Richardson, 1846), *Labeo coubie* Rüeppell, 1832 and *Acanthopsoides gracilentus* (Smith, 1945), that are all freshwater fishes from Southeast Asia or Africa and have never occurred in the sampling region, were used as internal standard DNAs. A target region (mitochondrial 12S rRNA) of the extracted DNA was amplified using MiFish primers (without MiSeq adaptors) (Miya et al. 2015) and the amplified and purified target DNA (about 220 bp) was excised using E-Gel SizeSelect (ThermoFisher Scientific, Waltham, MA, USA). The DNA size distribution of the library was estimated using an Agilent 2100 BioAnalyzer (Agilent, Santa Clara, CA, USA) and the concentration of double-stranded DNA of the library was quantified using a Qubit dsDNA HS assay kit and a Qubit fluorometer (ThermoFisher Scientific, Waltham, MA, USA). Based on the quantification values obtained using the Qubit fluorometer, the copy number of the standard DNAs was adjusted and these DNAs were mixed as follows: *S. immaculatus* (500 copies/µl), *E. bambusa* (250 copies/µl), *C. acuminatus* (100 copies/µl), *L. coubie* (50 copies/µl) and *A. gracilentus* (25 copies/µl). Hereafter, the mixed standard DNA is referred to as 'standard DNA mix'. The numbers of internal standard DNA copies added to samples were determined by quantification of the number of total fish eDNA copies (i.e. MiFish primer target region; Miya et al. 2015) using the SYBR-GREEN quantitative PCR method (see below for the detailed method).

Paired-end library preparation

Work-spaces and equipment were sterilised prior to the library preparation, filtered pipette tips were used and separate rooms were used for pre- and post-PCR operations to safeguard against cross-contamination. Negative controls were also employed to monitor contamination during the experiments. A fish universal primer set (MiFish primers; Miya et al. 2015) was used to amplify fish eDNA in the samples.

The first-round PCR (1st PCR) was carried out with a 12-µl reaction volume containing 6.0 µl of 2 × KAPA HiFi HotStart ReadyMix (KAPA Biosystems, Wilmington, WA, USA), 0.7 µl of each primer (5 µM), 0.6 µl of sterilised distilled H_2O, 2 µl of standard DNA mix and 2.0 µl of template. Note that the standard DNA mix was included for each sample. The final concentration of each primer was 0.3 µM. The sequences of MiFish primers are: GTCGGTAAAACTCGTGCCAGC (MiFish-U-F) and CATAGTGGGGTATCTAATCCCAGTTTG (MiFish-U-R) (Miya et al. 2015). MiSeq sequencing primers and six random bases (N) were combined with MiFish-U primers (see Miya et al. 2015 and Suppl. material 2 for detailed sequences). The six random bases were used to enhance cluster separation on the flowcells during initial base call calibrations on the MiSeq platform. The thermal cycle profile after an initial 3 min denaturation at 95°C was as follows (35 cycles): denaturation at 98°C for 20 s; annealing at 65°C for 15 s; and extension at 72°C for 15 s, with a final extension at the same temperature for 5 min. Triplicate 1st PCR were performed and the replicates were pooled in order to mitigate the PCR dropouts. Each pooled 1st PCR product (i.e. one pooled 1st PCR product per sample) was purified using Exo-SAPIT (Affymetrix, Santa Clara, CA, USA). The pooled, purified and 10-

fold diluted 1st PCR products were used as templates for the second-round PCR.

The second-round PCR (2nd PCR) was carried out with a 24-µl reaction volume containing 12 µl of 2 × KAPA HiFi HotStart ReadyMix, 1.4 µl of each primer (5 µM), 7.2 µl of sterilised distilled H2O and 2.0 µl of template. Different combinations of forward and reverse indices were used for different templates (samples) for massively parallel sequencing with MiSeq (Suppl. material 2). The thermal cycle profile after an initial 3 min denaturation at 95˚C was as follows (12 cycles): denaturation at 98˚C for 20 s; combined annealing and extension at 72˚C (shuttle PCR) for 15 s, with a final extension at 72˚C for 5 min. The products of the second PCR were combined (i.e. one pooled 2nd PCR product that included all samples), purified (using AMPure XP; PCR product:AMPure XP beads = 1:0.8; Beckman Coulter, Brea, California, USA), excised (using E-Gel SizeSelect; ThermoFisher Scientific, Waltham, MA, USA) and sequenced on the MiSeq platform using a MiSeq v2 Reagent Nano Kit for 2 × 150 bp PE with 5% PhiX spike-in.

Sequence read processing and taxonomic assignment

The detailed information about the above bioinformatics pipeline from data pre-processing through taxonomic assignment is available in the supplemental information in Miya et al. (2015). An online version of this pipeline is also available at http://mitofish.aori.u- tokyo.ac.jp/mifish.

The overall quality of the MiSeq reads was evaluated using FastQC (available from http:// www.bioinformatics.babraham.ac.uk/projects/fastqc/) and the reads were assembled using the software FLASH with a minimum overlap of 10 bp (Magoc and Salzberg 2011). The assembled reads were further filtered and cleaned and the pre-processed reads were subjected to the clustering process and taxonomic assignments. The pre-processed reads from the above custom pipeline were de-replicated using UCLUST (Edgar 2010). Those sequences represented by at least 10 identical reads were subjected to the downstream analyses and the remaining under-represented sequences (with less than 10 identical reads) were subjected to pairwise alignment using UCLUST. If the latter sequences (observed from less than 10 reads) showed at least 99% identity with one of the former reads (i.e. no more than one or two nucleotide differences), they were operationally considered as identical (with the differences being attributed to sequencing or PCR errors and/or actual nucleotide variations in the populations).

The processed reads were subjected to local BLASTN searches against a custom-made database (Camacho et al. 2009). The custom-made database was generated as described in a previous study (Miya et al. 2015). The database contains more than 7,000 fish species (http:// mitofish.aori.u-tokyo.ac.jp/; Iwasaki et al. 2013), which covers over 95% of fish species found in the study area. The top BLAST hit with a sequence identity of at least 97% and *E*-value threshold of 10^{-5} was applied for species assignments of each representative sequence.

Determination of the number of eDNA copies by quantitative PCR

The copy numbers of total fish eDNA were quantified using the SYBR-GREEN qPCR method using a StepOne-Plus™ Real-Time PCR system (Applied Biosystems, Foster City, CA, USA). SYBR-GREEN qPCR was conducted in a 10-µL volume with a reaction solution that consisted of 5 µl of PowerUpTM SYBR® GREEN Master Mix (Thermo Fisher Scientific, Wilmington, DE, USA), 0.6 µl of 5 µM MiFish-U-F/R primers (without adaptor), 10 µl of sterilised H2O and 1.0 µl of DNA template. SYBR-GREEN qPCR was performed in triplicate for each eDNA sample, the standard dilution series and PCR negative controls. The standard dilution series was prepared using DNA extracted from *Capoeta capoeta* (Guldenstadt, 1773) (a freshwater fish species in Southeast Asia). *C. capoeta*as was selected as the standard because the length of the MiFish region of this species is close to the average length in fish species (*C. capoeta* = 174 bp, the average length of the MiFish region = 173 bp). The thermal cycle profile after preconditioning for 2 min at 50˚C and 2 min at 95˚C was as follows (40 cycles): denaturation at 95˚C for 3 s; annealing and extension combined at 60˚C (shuttle PCR) for 30 s. Although MiFish primers predominantly amplify fish (e)DNA, it should be noted that the quantification by SYBR-GREEN qPCR may include non-fish eDNA because non-target sequences (e.g. sequences longer than the MiFish region) are sometimes amplified when using MiFish primers and because SYBR- GREEN qPCR does not distinguish between fish and non-fish eDNA. However, if the ratio of non-fish and fish amplicons does not drastically differ amongst samples, the SYBR- GREEN qPCR should reflect the dynamics (i.e. temporal fluctuation pattern) of total fish eDNA reasonably well. In all experiments, PCR negative controls showed no detectable amplification.

In addition, fish-species-specific eDNA was quantified by real-time TaqMan PCR according to using a StepOne-Plus™ Real-Time PCR system. The cytochrome *b* region of mitochondrial DNA was targeted for amplification from eDNA samples for each target species by using the following primer sets and associated probes, which were designed and confirmed as being able to amplify each target species specifically (Yamamoto et al. 2016 and Suppl. material 3). For the TaqMan qPCR analysis, *Engraulis japonicus* Temminck & Schlegel, 1846 (Japanese anchovy) and *Trachurus japonicus* (Temminck & Schlegel, 1844) (Japanese jack mackerel) were chosen because they are abundant in the study area and the standard dilution series were already available. For Japanese anchovy, primers Eja-CytB-Forward (5′-GAAAAACCCACCCCCTACTCA-3′), Eja-CytB-Reverse (5′-GTGGCCAAGCATAGTCCTAAAAG-3′) and Eja-CytB-Probe (5′-FAM- CGCAGTAGTAGACCTC-

CCAGCACCATCC-TAMRA-3′) were used. For Japanese jack mackerel, primers Tja-CytB-Forward (5′-CAGATATCGCAACCGCCTTT-3′), Tja-CytB- Reverse (5′-CCGATGTGAAGGTAAATGCAAA-3′) and Tja-CytB-Probe (5′-FAM- TATGCACGCCAACGGCG-CCT-TAMRA-3′) were used (summarised in Suppl. material 2). The length of the PCR amplicon produced using the primer set was 115 bp and 127 bp for Japanese anchovy and for Japanese jack mackerel, respectively. PCR was conducted in a 15-µl volume containing each primer at 900 nM, TaqMan® probe at 125 nM and 2 µl of sample DNA in 1 × PCR master mix (TaqMan® gene expression master mix; Life Technologies, Carlsbad, CA, USA). A dilution series of standards was prepared for quantification and analysed at the concentration of 3 × 101 to 3 × 104 copies per well in each experiment to obtain standard curves. The standards were pTAKN-2 plasmids containing commercially synthesised artificial DNA that had the same sequence as the amplification region of each species. The thermal cycle profile after preconditioning for 2 min at 50°C and 10 min at 95°C was as follows (55 cycles): denaturation at 95°C for 15 s; combined annealing and extension at 60°C (shuttle PCR) for 60 s. qPCR was performed in triplicate for each eDNA sample, standard dilution series and PCR negative controls. In all experiments, PCR negative controls showed no detectable amplification.

Statistical analysis

For all analyses, the free statistical environment R was used (R Core Team 2016). The statistical analyses consisted of three parts: linear regression analysis to examine:

1. the relationship between sequence reads and the copy numbers of the standard DNA for each sample,
2. the conversion of sequence reads of non-standard fish eDNA to calculated copy numbers using the result of the linear regression for each sample and
3. the comparison of eDNA copy numbers quantified by MiSeq and qPCR.

Linear regressions were performed using the *lm* function in R and used to examine how many sequence reads were generated from one (e)DNA copy through the library preparation process for MiSeq. Note that a linear regression between sequence reads and standard DNAs was performed for each sample and the intercept was set as zero. The regression equation was: MiSeq sequence reads = regression slope × the number of standard DNA copies [/µl]. The number of linear regressions performed was 52 (= the number of eDNA samples) and thus 52 regression slopes were estimated in total (see Suppl. material 4).

The sequence reads of non-standard fish eDNAs were converted to calculated copy numbers using a sample-specific regression slope estimated by the first analysis. The number of non-standard eDNA copies was estimated by dividing MiSeq sequence reads by a sample-specific regression slope (i.e. the number of DNA copies = MiSeq sequence reads/regression slope; hereafter, this equation is referred to as 'correction equation'). The estimated numbers of non-standard fish eDNA copies are hereafter referred to as 'calculated copy numbers' and this method itself (i.e. from an inclusion of standard DNA to the conversion of sequence reads using a correction equation) is hereafter referred to as 'qMiSeq'.

Calculated copy numbers by qMiSeq were compared with copy numbers estimated by qPCR (see above sections for detailed qPCR method) by using four approaches. First, raw values (non-standardised copy numbers) were compared using linear regressions (i.e. first approach). As there were significant outliers, linear regressions were again performed by excluding the outliers (i.e. second approach). In addition, because the distribution of calculated copy numbers was highly right-skewed (i.e. many samples with low copy numbers and few samples with high copy numbers), log-transformation (base = 2) was applied after adding 0.5 to the raw values. The log-transformed values were further compared using linear regressions (i.e. third approach). Lastly, a Bland-Altman plot (difference plot) (Bland and Altman 1986) was constructed (for raw and log-transformed values) to measure the agreement between qMiSeq and qPCR using BlandAltmanLeh package (Lehnert 2015) (i.e. fourth approach). Linear relationships were considered significant if *P* values were less than 0.05.

Data and code availability

Sequences are deposited in DDBJ Sequence Read Archive (DRA): Accession numbers are: DRA005598 (Submission ID), PRJDB5570 (BioProject ID) and SAMD00075651– SAMD00075720 (BioSample ID). All R codes and original data table used for the analyses are available at https://doi.org/10.5281/zenodo.1173828.

Results and discussion

Relationship between the copy numbers and sequence reads of the standard DNA

The sequence reads of the internal standard DNAs were significantly positively correlated with the copy numbers of those DNAs (Fig. 2a and b, Suppl. material 4). A regression line was drawn for each eDNA sample and therefore the number of regression lines equalled the number of eDNA samples (= 52; Suppl. material 4). R^2 values of the regression lines ranged from 0.71 to 0.98 and more than 80% of regression lines showed R^2 values higher than 0.9 (Fig. 2c; see Suppl. material 5 for regression residuals), suggesting that the number of sequence reads was proportional to the number of DNA copies in a single sample and that the slopes of the regression lines (i.e.sequence reads per DNA copy) can be used to convert sequence reads to the numbers of DNA copies.

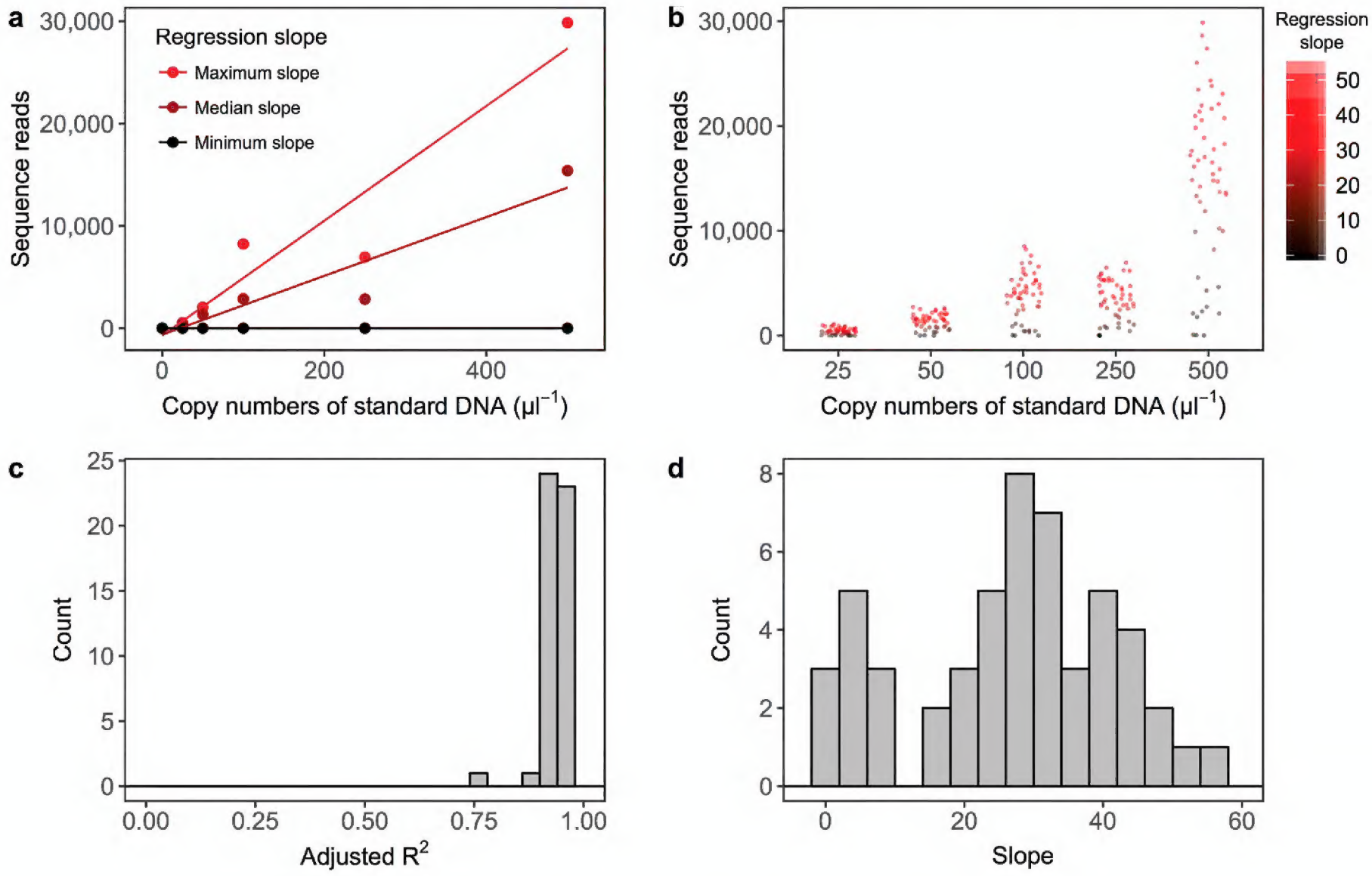

Figure 2. Summary of regression lines constructed using the number of copies added and sequence reads of internal standard DNAs. Examples of a regression line (**a**). Regression lines with the maximum, median and minimum slopes are indicated as examples of the relationships. The line indicates the regression line between the copy numbers of standard DNA (copies/µl) and sequence reads. The intercept of the regression line was set as zero. Distributions of sequence reads of internal standard DNAs (**b**). The intensity of red colour indicates the slope of regression line. Distribution of adjusted R^2 of the regression lines (**c**). Note that a regression line was drawn for each eDNA sample and that the number of standard curves is equal to the number of eDNA samples ($N = 52$). Distribution of slopes of regression lines (**d**).

Interestingly, the slopes of the regression lines were highly variable, ranging from 0 to 54.1 (which corresponded to eDNA samples collected on 16 June 2015 and 21 April 2015, respectively), with a median value of 24.6 (Fig. 2d, Suppl. material 4). Low slope values (e.g. 0 or close to 0) indicate that internal standard DNAs were not efficiently amplified even if the number of DNA copies added was large, suggesting the presence of PCR inhibitor(s) (e.g. humic substance) in the eDNA samples. These variations of the slope also suggested that the degree of PCR inhibition varies depending on the eDNA sample.

Quantification of the copy number using sequence reads and correction equations and comparison of the calculated copy number with the copy number quantified by qPCR

MiSeq sequence reads of each sample were converted using each correction equation (i.e. the number of eDNA copies [copies/µl] = MiSeq sequence reads / a sample-specific regression slope; the copy numbers and this method itself are referred to as 'calculated DNA copies [copies/µl]' and 'qMiSeq', respectively). Then, calculated DNA copies were compared with the number of DNA copies quantified by qPCR (Fig. 3 [all regression lines were significant, $P < 0.05$] and Suppl. materials 6, 7). The numbers of eDNA copies estimated by qMiSeq and qPCR were significantly and positively correlated with each other for total fish eDNA (all data included, Fig. 3a; outliers excluded, Fig. 3b). For the total fish eDNA, the number of eDNA copies quantified by qPCR (mean copy number = 683 copies/µl) was higher than that quantified by qMiSeq (mean copy number = 139 copies/µl). This is not surprising because target amplicon fragments (about 370 bp, including MiSeq adaptor) were excised and non-target amplified fragments (e.g. longer and unknown amplicons), which were included in the quantification by the SYBR-GREEN assay, were discarded before MiSeq sequencing. Regarding the eDNA of Japanese anchovy and Japanese jack mackerel, it was found that the number of eDNA copies quantified by qMiSeq was similar to that obtained by qPCR (i.e. regression lines were close to the 1:1 line in Fig. 3c–f, regardless of the inclusion/exclusion of outliers; for the relationships between sequence reads and copy numbers quantified by qPCR, see Suppl. material 8). In addition, the Bland-Altman plot for the raw

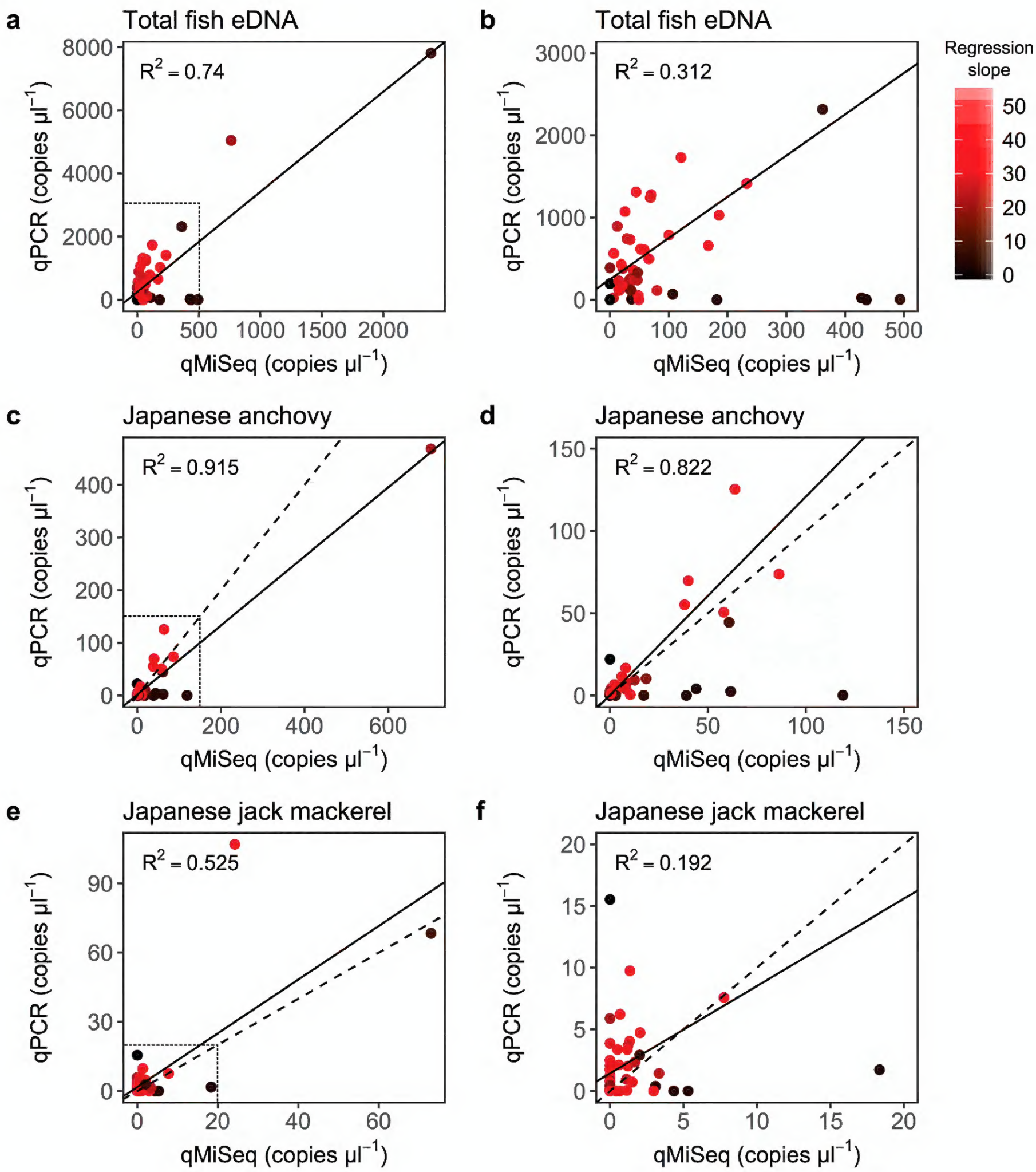

Figure 3. Relationship between the number of eDNA copies quantified by qPCR and that by qMiSeq. Correlations for the total fish eDNA (all data, **a**; enlarged figure, **b**), Japanese anchovy (*Engraulis japonicus*; all data, **c**; enlarged figure, **d**) and Japanese jack mackerel (*Trachurus japonicus*; all data, **e**; enlarged figure, **f**). Dashed and solid lines indicate 1:1 line and linear regression line, respectively. Regression lines in the enlarged figures were drawn by excluding outliers. All regression lines were significant ($P < 0.05$). Dotted boxed regions in **a**, **c** and **e** correspond to the range of the graphs in **b**, **d** and **f**, respectively. The intensity of red colour indicates the slope of the regression line (=correction equation) used to convert sequence reads to the calculated copy numbers.

values also showed that the differences in the copy numbers quantified by the two methods were not significantly different from zero (Suppl. material 9). These results suggested that eDNA metabarcoding with the inclusion of internal standard DNAs reasonably quantified the number of eDNA copies.

Although the calculated DNA copies generally corresponded well with the eDNA copy numbers estimated by qPCR, the calculated DNA copies of some samples were

much higher than the copy numbers obtained by qPCR (i.e. for the points close to the *x*-axis in Fig. 3). These samples showed relatively small values for the slopes of regression lines between the sequence reads and quantity of the standard DNAs (i.e. corresponded to points with darker colour in Fig. 3), suggesting that there was inhibition of PCR in these samples (PCR inhibitions often occur in environmental samples; Schrader et al. 2012). qMiSeq can control PCR-inhibition effects in the estimation of eDNA copy because correction equations already take the influence of PCR inhibition into account, which may be an advantage of this method compared with qPCR. Conversely, it is suggested that qPCR could not reliably quantify the number of eDNA copies when the influence of PCR inhibitors was strong. Theoretically, influences of PCR inhibition could also be tested by (multiplex) qPCR (e.g. Hartman et al. 2005, Turner et al. 2015), but multiplex qPCR may require some additional experimental procedures and thus could be more time-consuming and costly.

Some samples showed much lower eDNA copy numbers of Japanese jack mackerel when quantified by qMiSeq than when quantified by qPCR (i.e. points close to the *y*-axis; Fig. 3f). This inconsistency might have been due to the low eDNA copy number of Japanese jack mackerel (all samples showed less than 100 copies/µl and most samples showed less than 10 copies/µl). qMiSeq might not be able to quantify such low numbers of eDNA copies accurately because the lowest copy number of internal standard DNA added was 25 copies/µl. If the copy number of internal standard DNA had been much lower, more accurate quantification would have been achieved by qMiSeq. Furthermore, the difference in number of PCR cycles between qPCR (40–55 cycles) and the 1st PCR of MiSeq library preparation (35 cycles) might contribute to the different sensitivities (i.e. detection limits) of these methods.

Although the above analyses suggested that there is good agreement between the two methods, the distributions of calculated copy numbers as well as copy numbers estimated by qPCR were right-skewed (i.e. many low copy numbers and few high copy numbers) and thus the copy numbers were further compared after log-transformation of the raw values. Samples with regression slopes lower than 10 were excluded from this analysis because they suggested that there had been significant PCR inhibition during the qPCR measurements, as discussed above. It was found that there were positive and significant linear relationships between the copy numbers quantified by qMiSeq and qPCR even after the log-transformation ($P < 0.05$; Fig. 4a–c). Bland-Altman plots also suggested that there is a good agreement between the two methods (i.e. 95% confidence intervals include zero; Fig. 4e and f). Taken together, these results suggested that eDNA metabarcoding with internal standard DNA enabled simultaneous quantification and identification of fish eDNA. The appropriate range of the copy numbers of the internal standard DNAs should, however, be carefully determined depending on the range of the target eDNA copy numbers in environmental samples.

Temporal dynamics of Maizuru-bay fish eDNA revealed by qMiSeq

The temporal dynamics of the fish eDNA quantified by qMiSeq generally corresponded well with those quantified by qPCR (Fig. 5). For the total fish eDNA, the highest eDNA concentrations were found on 25th August and 24th November by qPCR and peaks were also detected on those dates by qMiSeq (Fig. 5a). For Japanese anchovy eDNA, the highest eDNA concentration was found on 24th November by qPCR and the peak was also detected on this date by qMiSeq (Fig. 5b). For Japanese jack mackerel eDNA, one of the highest eDNA concentrations (on 25th August) found by qPCR was also found by qMiSeq. However, another peak found by qPCR (on 23rd February) was not detected by qMiSeq (Fig. 5c), probably due to the above-mentioned technical issues in eDNA metabarcoding with internal standard DNA.

The results obtained in the present study suggest that qMiSeq can reasonably recover the dynamics of fish eDNA. As eDNA metabarcoding can detect many species (sometimes more than 100 species) in a single run (Miya et al. 2015), this method enables simultaneous quantifications of eDNA derived from manyfish species. In the present study, more than 70 fish species were detected from 52 eDNA samples collected from April 2015 to March 2016 in Maizuru Bay, Japan (Suppl. materials 3, 6), which is generally consistent with long-term direct visual observations, e.g. fortnightly-performed visual census over 5 years detected a total of 83 fish species (Masuda 2008, Masuda et al. 2010).

Quantitative and multispecies fish eDNA monitoring

This method enables the generation of a quantitative time series of eDNA for these fish species by a single MiSeq run and, as an example, an eDNA time series of the 10 most abundant fish species in terms of eDNA concentration is shown in Fig. 6. As eDNA copy numbers may be a rough index for fish biomass/abundance (Takahara et al. 2012), such a multispecies quantitative time series, which can readily be obtained if the qMiSeq method is used, may provide valuable information about the dynamics of fish populations in the sampling area. This method did not correct (fish) species-specific PCR amplification biases (see Elbrecht and Leese (2015), Krehenwinkel et al. (2017) for taxon-specific PCR biases), but the eDNA time series measured here by qMiSeq were ecologically interpretable, suggesting that eDNA monitoring using this method would provide ecologically meaningful information on the dynamics of a natural fish community, at least in this case (see Suppl. materials 3, 10).

As eDNA metabarcoding has been recognised as an efficient approach in species detection and biodiversity assessment (Fukumoto et al. 2015, Miya et al. 2015, Sigsgaard et al. 2015, Deiner et al. 2016, Ushio et al. 2017a, b, Yamamoto et al. 2017), its use as a biodiversity monitoring tool has been increasing (Sigsgaard et al. 2015, Bista

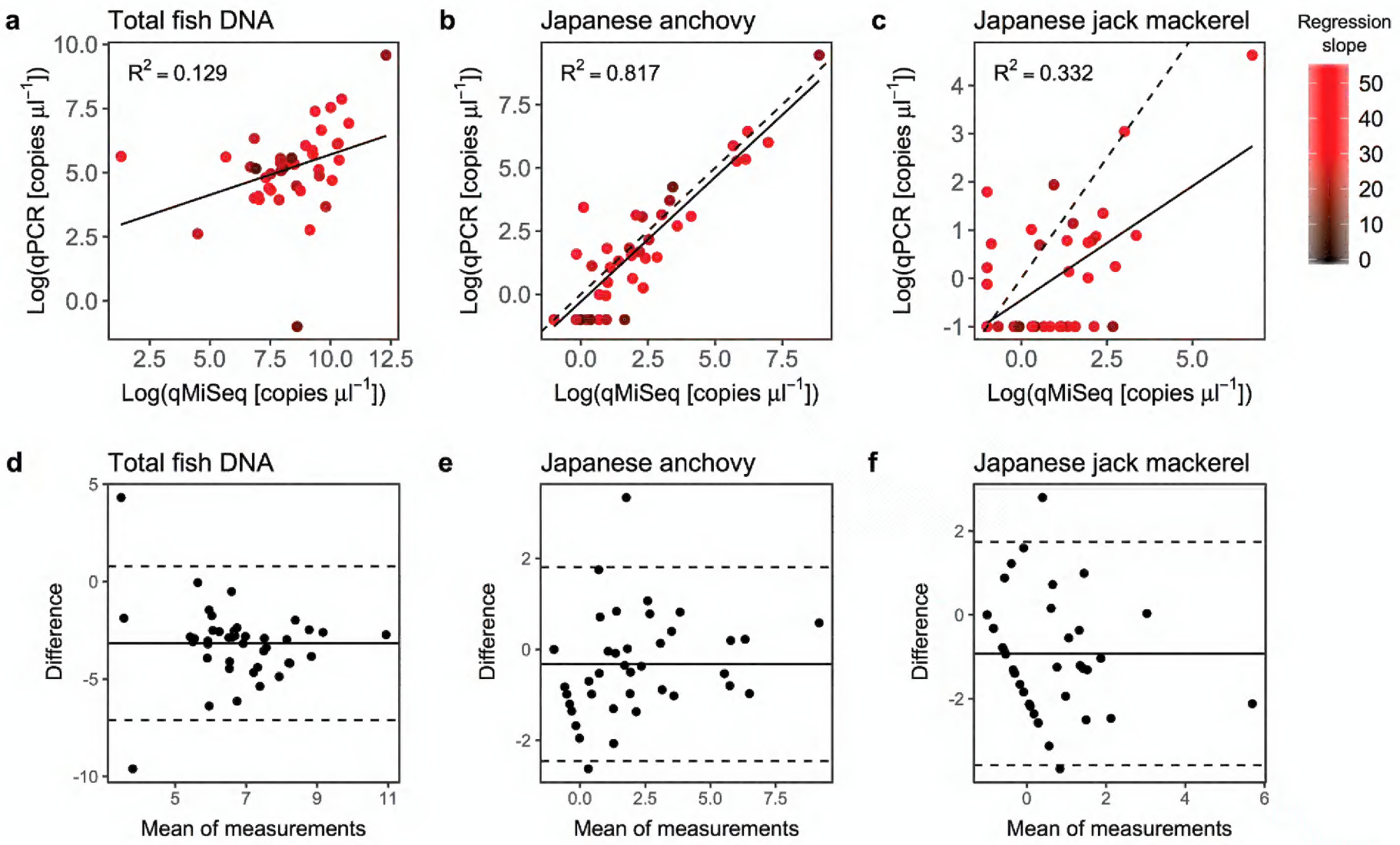

Figure 4. Relationships between log-transformed copy numbers quantified by qPCR and qMiSeq for the total fish eDNA (**a**), Japanese anchovy (*Engraulis japonicus*; **b**) and Japanese jack mackerel (*Trachurus japonicus*; **c**). A solid line is a linear regression line (all lines are significant; $P < 0.05$). The intensity of red colour indicates the slope of the regression line (i.e. correction equation) used to convert sequence reads to the copy numbers. Bland-Altman plot for log- transformed copy numbers of the total fish eDNA (**d**), Japanese anchovy (**e**) and Japanese jack mackerel (**f**). Dashed lines indicate 95% upper and lower limits and solid line indicates mean values. Note that, although it is not significant, the Bland-Altman plot for the log- transformed copy numbers of the total fish eDNA (**d**) showed that the calculated copy numbers by qMiSeq tends to be smaller than those by qPCR probably because of the removal of non-target amplified fragments before MiSeq sequencing (see discussion in the main text).

et al. 2017, Stoeckle et al. 2017). Those monitoring studies performed periodic water samplings and generated eDNA time series and showed that temporal fluctuations in species (or OTU) richness and detection probability of eDNA of a target taxa were in good agreement with temporal fluctuations in other reliable data (e.g. visual census) (Sigsgaard et al. 2015, Hänfling et al. 2016, Bista et al. 2017, Stoeckle et al. 2017) . However, because of a lack of a quantitative method for evaluating eDNA metabarcoding, only qualitative information about eDNA (e.g. presence/absence, rank of eDNA sequence reads and species/OTU diversity) has been reported for the comparisons between eDNA monitoring data and other monitoring data. The use of sequence reads as a quantitative index of the abundance/biomass of target organisms may partly solve this problem (Evans et al. 2015). However, the number of sequence reads per sample (or per species) may change dramatically depending on experimental conditions such as the number of samples multiplexed, final library concentrations and sequence reagents and thus rigorous comparisons between samples originated from different experiments/studies are difficult.

The quantities of internal standards are precisely known and thus the use of an internal standard would enable rigorous and quantitative comparisons even between different experiments/studies (given the same PCR primers are used), which would facilitate the use of eDNA metabarcoding as a tool for biodiversity monitoring. Furthermore, this method, i.e. the addition of purified DNA fragments, would be less time- and effort-consuming than the use of tissues of standard organisms as internal standards (Thomas et al. 2015, Smets et al. 2016) because the preparation of standard organisms/tissues is sometimes difficult. In future studies, the use of artificial fish sequences, that are not identical to the sequences of any other fish species in the world, should be considered because it would be applicable to any water sample and would further increase the efficiency of this method.

Conclusions

In the present study, it was shown that eDNA metabarcoding, performed with the inclusion of internal standard DNA, enables simultaneous determination of the quantity and identity of eDNA derived from multiple fish species. As the traditional species-specific qPCR allows

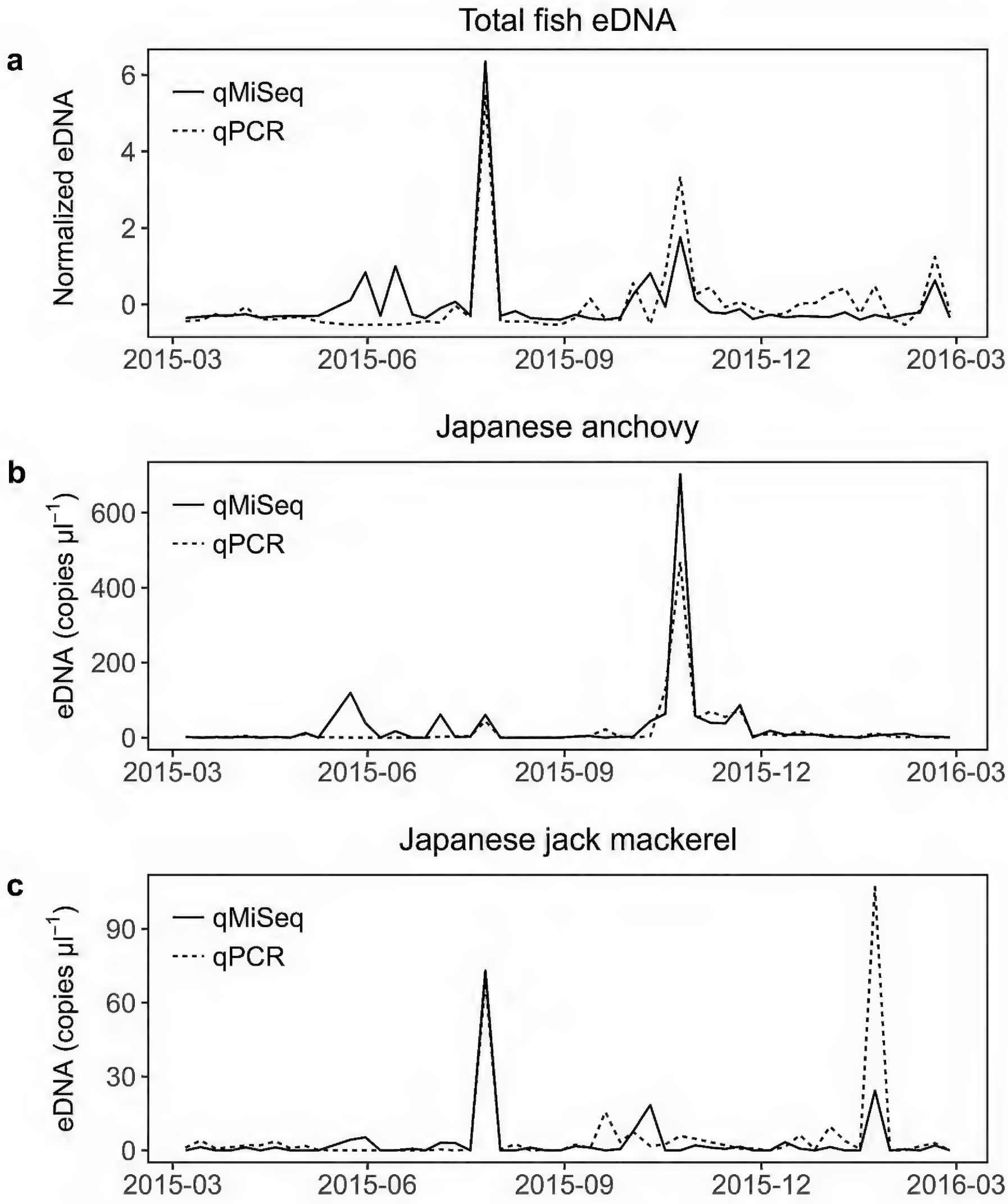

Figure 5. Dynamics of the total fish eDNA (**a**), Japanese anchovy (*Engraulis japonicus*; **b**) and Japanese jack mackerel (*Trachurus japonicus*; **c**) quantified by qMiSeq and qPCR. Solid and dashed lines indicate the number of eDNA copies quantified by qMiSeq and qPCR, respectively. Note that the copy numbers of total fish eDNA were normalised to have zero mean and unit variance.

quantification of eDNA from only one fish species in a single experiment, this method is much more efficient compared with qPCR. In addition, this method can take effects of PCR inhibition into account. Although it should be mentioned that fish eDNA copy numbers are still only a rough index of fish biomass/abundance (or population size) and this problem should be addressed in a future study (e.g. estimating taxon-specific correction factors is a promising direction; see Krehenwinkel et al. 2017), these results show that eDNA metabarcoding with the inclusion of internal standard DNAs can be a promising tool to monitor fish biodiversity. This method will improve the efficiency of obtaining data and may contribute to more effective resource management and ecosystem monitoring.

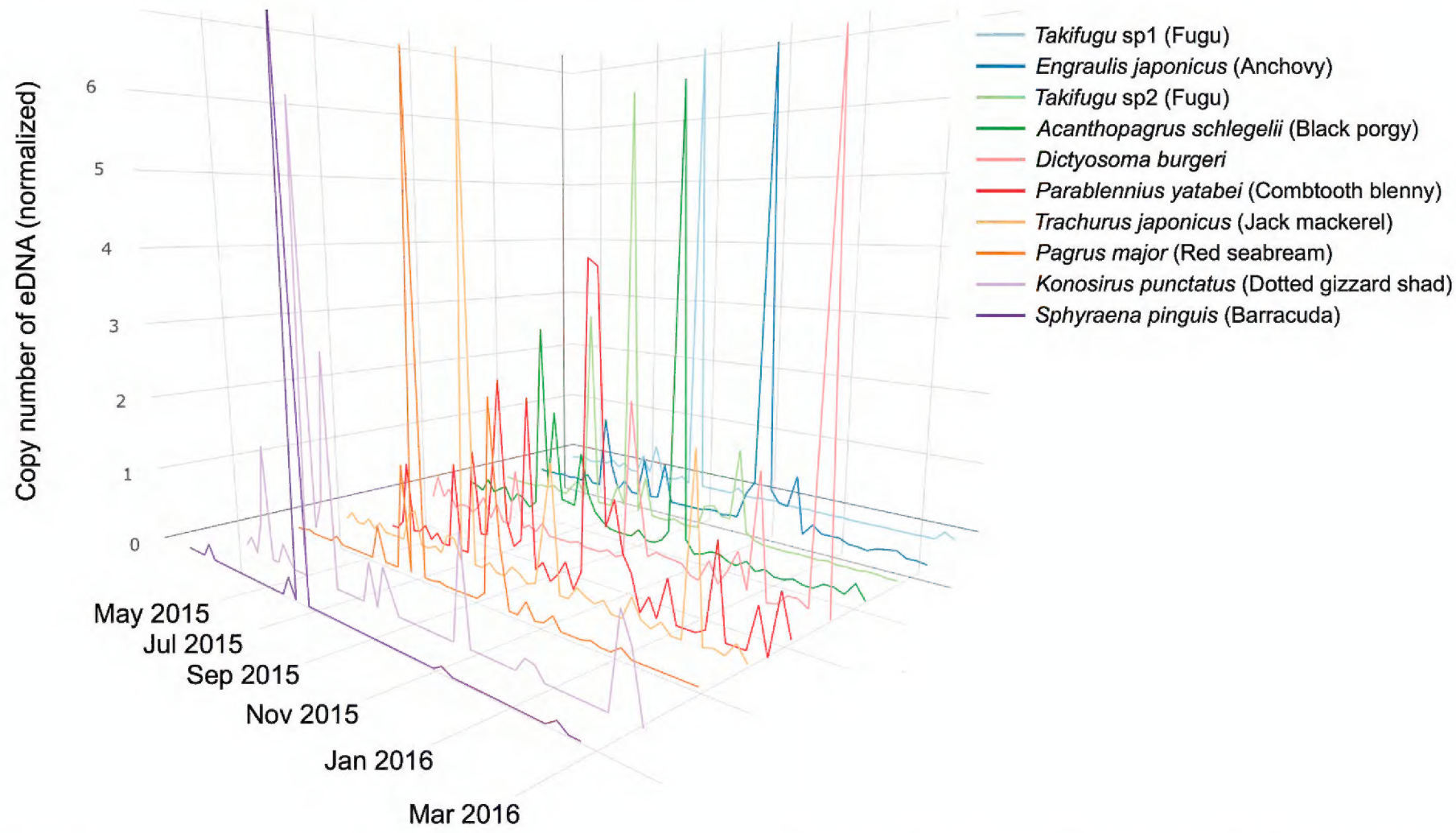

Figure 6. Quantitative and multispecies fish eDNA time series in Maizuru Bay, Japan. Time series of eDNA of 10 dominant fish species. In the eDNA analysis, two *Takifugu* species were detected as dominant species and were designated *Takifugu* sp1 and sp2. A representative sequence of *Takifugu* sp1 is highly similar to that of *T. niphobles*/*T. snyderi* (>99% identity). A representative sequence of *Takifugu* sp2 is identical with that of *T. pardadalis*/*T. xanthopterus*/ *T. poecilonotus* (100% identity). Different colours indicate different fish species. The numbers of eDNA copies were normalised to have zero mean and unit variance.

Funding program

This research was supported by CREST (JMPJ-CR13A2) from the Japan Science and Technology Agency (JST).

Ethics and security

The experiments were conducted in accordance with the guidelines of Regulation of Animal Experimentation at Kyoto University. Water sampling permission in or around the pier was not needed.

Author contributions

MU conceived and designed research; HM and RM performed sampling; MU, HM, RM, TS, MM, SS, HY and TM performed experiments; MU analysed data; MU wrote the early draft and completed it with significant inputs from all authors.

Conflicts of interest

The authors declare no competing interests.

Acknowledgements

We would like to thank Aina Tanimoto for help with the collection and filtration of water samples and Hideyuki Doi for providing meaningful comments on the manuscript. We also thank The Research Center for Satoyama Studies at Ryukoku University for providing an opportunity to use the Illumina MiSeq platform.

References

Barnes M, Turner C (2016) The ecology of environmental DNA and implications for conservation genetics. Conservation Genetics 17(1): 1–17. https://doi.org/10.1007s10592-015-0775-4

Baschien C, Manz W, Neu TR, Marvanova L, Szewzyk U (2008) *In situ* detection of freshwater fungi in an alpine stream by new taxon-specific fluorescence *in situ* hybridization probes. Applied and Environmental Microbiology 74(20): 6427–6436. https://doi.org/10.1128/aem.00815-08

Bista I, Carvalho G, Walsh K, Seymour M, Hajibabaei M, Lallias D, Christmas M, Creer S (2017) Annual time-series analysis of aqueous eDNA reveals ecologically relevant dynamics of lake ecosystem biodiversity. Nature Communications 8: 14087. https:/doi.org/10.1038/ncomms14087

Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet (London, England) 1(8476): 307–10. https:/doi.org/10.1016/S0140-6736(86)90837-8

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: architecture and applications. BMC Bioinformatics 10(1): 421. https:/doi.org/10.1186/1471-2105-10-421

Davy C, Kidd A, Wilson C (2015) Development and validation of environmental DNA (eDNA) markers for detection of freshwater turtles. PLoS ONE 10(7): e0130965. https://doi.org/10.1371/journal.pone.0130965

Deiner K, Fronhofer E, Mächler E, Walser J, Altermatt F (2016) Environmental DNA reveals that rivers are conveyer belts of biodiversity information. Nature Communications 7: 12544. https://doi.org/10.1038/ncomms12544

Dejean T, Valentini A, Duparc A, Pellier-Cuit S, Pompanon F, Taberlet P, Miaud C (2011) Persistence of environmental DNA in freshwater ecosystems. PLoS ONE 6(8): e23398. https://doi.org/10.1371/journal.pone.0023398

Edgar R (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26(19): 2460–2461. https://doi.org/10.1093/bioinformatics/btq461

Elbrecht V, Leese F (2015) Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass—sequence relationships with an innovative metabarcoding protocol. PLoS ONE 10(7): e0130324. https://doi.org/10.1371/journal.pone.0130324

Evans N, Olds B, Renshaw M, Turner C, Li Y, Jerde C, Mahon A, Pfrender M, Lamberti G, Lodge D (2015) Quantification of mesocosm fish and amphibian species diversity via environmental DNA metabarcoding. Molecular Ecology Resources 16(1): 29–41. https:/doi.org/10.1111/1755-0998.12433

Ficetola GF, Miaud C, Pompanon F, Taberlet P (2008) Species detection using environmental DNA from water samples. Biology Letters 4(4): 423–425. https:/doi.org/10.1098/rsbl.2008.0118

Fukumoto S, Ushimaru A, Minamoto T (2015) A basin-scale application of environmental DNA assessment for rare endemic species and closely related exotic species in rivers: a case study of giant salamanders in Japan. Journal of Applied Ecology 52(2): 358–365. https://doi.org/10.1111/1365-2664.12392

Hänfling B, Handley LL, Read D, Hahn C, Li J, Nichols P, Blackman R, Oliver A, Winfield I (2016) Environmental DNA metabarcoding of lake fish communities reflects long-term data from established survey methods. Molecular Ecology 25(13): 3101–3119. https://doi.org/10.1111/mec.13660

Hartman L, Coyne S, Norwood D (2005) Development of a novel internal positive control for Taqman® based assays. Molecular and Cellular Probes 19(1): 51–59. https://doi.org/10.1016/j.mcp.2004.07.006

Iwasaki W, Fukunaga T, Isagozawa R, Yamada K, Maeda Y, Satoh TP, Sado T, Mabuchi K, Takeshima H, Miya M, Nishida M (2013) MitoFish and MitoAnnotator: A mitochondrial genome database of fish with an accurate and automatic annotation pipeline. Molecular Biology and Evolution 30(11): 2531–2540. https://doi.org/10.1093/molbev/mst141

Jerde C, Mahon A, Chadderton WL, Lodge D (2011) "Sight-unseen" detection of rare aquatic species using environmental DNA. Conservation Letters 4(2): 150–157. https:/doi.org/10.1111/j.1755-263x.2010.00158.x

Kelly R, Port J, Yamahara K, Crowder L (2014) Using environmental DNA to census marine fishes in a large mesocosm. PLoS ONE 9(1): e86175. https://doi.org/10.1371journal.pone.0086175

Kelly RP, Port JA, Yamahara KM, Martone RG, Lowell N, Thomsen PF, Mach ME, Bennett M, Prahler E, Caldwell MR, Crowder LB (2014) Harnessing DNA to improve environmental management. Science 344(6191): 1455–1456. https://doi.org/10.1126science.1251156

Klobucar S, Rodgers T, Budy P (2017) At the forefront: evidence of the applicability of using environmental DNA to quantify the abundance of fish populations in natural lentic waters with additional sampling considerations. Canadian Journal of Fisheries and Aquatic Sciences 74(12): 2030–2034. https://doi.org/10.1139/cjfas-2017-0114

Krehenwinkel H, Wolf M, Lim JY, Rominger A, Simison W, Gillespie R (2017) Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding. Scientific Reports 7(1): #17668. https://doi.org/10.1038/s41598-017-17333-x

Lehnert B (2015) BlandAltmanLeh: Plots (Slightly Extended) Bland-Altman Plots. https://CRAN.R-project.org/package=BlandAltmanLeh [R package version 0.3.1]

Magoc T, Salzberg SL (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27(21): 2957–2963. https://doi.org/10.1093bioinformatics/btr507

Maruyama A, Nakamura K, Yamanaka H, Kondoh M, Minamoto T (2014) The release rate of environmental DNA from juvenile and adult fish. PLoS ONE 9(12): e114639. https://doi.org/10.1371/journal.pone.0114639

Masuda R (2008) Seasonal and interannual variation of subtidal fish assemblages in Wakasa Bay with reference to the warming trend in the Sea of Japan. Environmental Biology of Fishes 82(4): 387–399. https://doi.org/10.1007/s10641-007-9300-z

Masuda R, Shiba M, Yamashita Y, Ueno M, Kai Y, Nakanishi A, Torikoshi M, Tanaka M (2010) Fish assemblages associated with three types of artificial reefs: density of assemblages and possible impacts on adjacent fish abundance. Fishery Bulletin 108(2): 162–173.

Minamoto T, Yamanaka H, Takahara T, Honjo M, Kawabata Z (2012) Surveillance of fish species composition using environmental DNA. Limnology 13 (2): 193–197. https:/doi.org/10.1007/s10201-011-0362-4

Miya M, Sato Y, Fukunaga T, Sado T, Poulsen JY, Sato K, Minamoto T, Yamamoto S, Yamanaka H, Araki H, Kondoh M, Iwasaki W (2015) MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: detection of more than 230 subtropical marine species. Royal Society Open Science 2(7): 150088. https:/doi.org/10.1098/rsos.150088

R Core Team (2016) R: A language and environment for statistical computing. 3.3.0. https://www.r-project.org/

Samejima H, Ong R, Lagan P, Kitayama K (2012) Camera-trapping rates of mammals and birds in a Bornean tropical rainforest under sustainable forest management. Forest Ecology and Management 270: 248–256. https://doi.org/10.1016/j.foreco.2012.01.013

Schrader C, Schielke A, Ellerbroek L, Johne R (2012) PCR inhibitors – occurrence, properties and removal. Journal of Applied Microbiology 113(5): 1014–1026. https:/doi.org/10.1111/j.1365-2672.2012.05384.x

Sigsgaard EE, Carl H, Møller PR, Thomsen PF (2015) Monitoring the near-extinct European weather loach in Denmark based on environmental DNA from water samples. Biological Conservation 183: 46–52. https://doi.org/10.1016/j.biocon.2014.11.023

Simmons M, Tucker A, Chadderton WL, Jerde C, Mahon A (2016) Active and passive environmental DNA surveillance of aquatic invasive species. Canadian Journal of Fisheries and Aquatic Sciences 73(1): 76–83. https://doi.org/10.1139/cjfas-2015-0262

Smets W, Leff J, Bradford M, McCulley R, Lebeer S, Fierer N (2016) A method for simultaneous measurement of soil bacterial abundances

and community composition via 16S rRNA gene sequencing. Soil Biology and Biochemistry 96: 145–151. https:/doi.org/10.1016/j.soilbio.2016.02.003

Stoeckle M, Soboleva L, Charlop-Powers Z (2017) Aquatic environmental DNA detects seasonal fish abundance and habitat preference in an urban estuary. PLoS ONE 12(4): e0175186. https://doi.org/10.1371/journal.pone.0175186

Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E (2012) Towards next- generation biodiversity assessment using DNA metabarcoding. Molecular Ecology 21(8): 2045–2050. https://doi.org/10.1111/j.1365-294x.2012.05470.x

Takahara T, Minamoto T, Yamanaka H, Doi H, Kawabata Z (2012) Estimation of fish biomass using environmental DNA. PLoS ONE 7(4): e35868. https://doi.org/10.1371journal.pone.0035868

Takahara T, Minamoto T, Doi H (2013) Using environmental DNA to estimate the distribution of an invasive fish species in ponds. PLoS ONE 8(2): e56584. https:/doi.org/10.1371/journal.pone.0056584

Thomas A, Deagle B, Eveson JP, Harsch C, Trites A (2015) Quantitative DNA metabarcoding: improved estimates of species proportional biomass using correction factors derived from control material. Molecular Ecology Resources 16(3): 714–726. https://doi.org/10.1111/1755-0998.12490

Thomsen PF, Kielgast J, Iversen L, Wiuf C, Rasmussen M, P Gilvert MT, Orlando L, Willerslev E (2011) Monitoring endangered freshwater biodiversity using environmental DNA. Molecular Ecology 21(11): 2565–2573. https://doi.org/10.1111/j.1365-294x.2011.05418.x

Tsuji S, Ushio M, Sakurai S, Minamoto T, Yamanaka H (2017) Water temperature- dependent degradation of environmental DNA and its relation to bacterial abundance. PLoS ONE 12(4): e0176608. https://doi.org/10.1371/journal.pone.0176608

Turner C, Uy K, Everhart R (2015) Fish environmental DNA is more concentrated in aquatic sediments than surface water. Biological Conservation 183: 93–102. https:/doi.org/10.1016/j.biocon.2014.11.017

Ushio M, Murata K, Sado T, Nishiumi I, Takeshita M, Iwasaki W, Miya M (2017a) Demonstration of the potential of environmental DNA as a tool for the detection of avian species. bioRxi. https://doi.org/10.1101/199208

Ushio M, Fukuda H, Inoue T, Makoto K, Kishida O, Sato K, Murata K, Nikaido M, Sado T, Sato Y, Takeshita M, Iwasaki W, Yamanaka H, Kondoh M, Miya M (2017b) Environmental DNA enables detection of terrestrial mammals from forest pond water. Molecular Ecology Resources 17(6): 63–75. https://doi.org/10.1111/1755-0998.12690

Yamamoto S, Minami K, Fukaya K, Takahashi K, Sawada H, Murakami H, Tsuji S, Hashizume H, Kubonaga S, Horiuchi T, Hongo M, Nishida J, Okugawa Y, Fujiwara A, Fukuda M, Hidaka S, Suzuki K, Miya M, Araki H, Yamanaka H, Maruyama A, Miyashita K, Masuda R, Minamoto T, Kondoh M (2016) Environmental DNA as a 'Snapshot' of fish distribution: A case study of Japanese jack mackerel in Maizuru Bay, Sea of Japan. PLoS ONE 11(3): e0149786. https://doi.org/10.1371/journal.pone.0149786

Yamamoto S, Masuda R, Sato Y, Sado T, Araki H, Kondoh M, Minamoto T, Miya M (2017) Environmental DNA metabarcoding reveals local fish communities in a species-rich coastal sea. Scientific Reports 7: 40368. https://doi.org/10.1038/srep40368

Supplementary material 1

The numbers of sequence reads remaining (filtered) in data processing steps

Authors: Masayuki Ushio, Hiroaki Murakami, Reiji Masuda, Tetsuya Sado, Masaki Miya, Sho Sakurai, Hiroki Yamanaka, Toshifumi Minamoto, Michio Kondoh

Data type: Supplementary Table

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/mbmg.2.23297.suppl1

Supplementary material 2

Primer, index and probe sequences used in the study

Authors: Masayuki Ushio, Hiroaki Murakami, Reiji Masuda, Tetsuya Sado, Masaki Miya, Sho Sakurai, Hiroki Yamanaka, Toshifumi Minamoto, Michio Kondoh

Data type: Supplementary Table

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/mbmg.2.23297.suppl2

Supplementary material 3

Descriptions of TaqMan probe specificity test and supplementary Table S1 and S2

Authors: Masayuki Ushio, Hiroaki Murakami, Reiji Masuda, Tetsuya Sado, Masaki Miya, Sho Sakurai, Hiroki Yamanaka, Toshifumi Minamoto, Michio Kondoh

Data type: Supplementary Text

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/mbmg.2.23297.suppl3

Supplementary material 4

The relationship between MiSeq sequence reads and copy numbers of standard DNAs for 52 samples

Authors: Masayuki Ushio, Hiroaki Murakami, Reiji Masuda, Tetsuya Sado, Masaki Miya, Sho Sakurai, Hiroki Yamanaka, Toshifumi Minamoto, Michio Kondoh

Data type: Supplementary Figure

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The

Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/mbmg.2.23297.suppl4

Supplementary material 5

The relationship between regression residuals and copy numbers of standard DNAs for 52 samples

Authors: Masayuki Ushio, Hiroaki Murakami, Reiji Masuda, Tetsuya Sado, Masaki Miya, Sho Sakurai, Hiroki Yamanaka, Toshifumi Minamoto, Michio Kondoh

Data type: Supplementary Figure

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/mbmg.2.23297.suppl5

Supplementary material 6

The numbers of eDNA copies of marine fish species quantified by metabarcoding with the internal standard DNA

Authors: Masayuki Ushio, Hiroaki Murakami, Reiji Masuda, Tetsuya Sado, Masaki Miya, Sho Sakurai, Hiroki Yamanaka, Toshifumi Minamoto, Michio Kondoh

Data type: Supplementary Table

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/mbmg.2.23297.suppl6

Supplementary material 7

Results of quantitative PCR for total fish eDNA, Japanese anchovy and Japanese jack mackerel

Authors: Masayuki Ushio, Hiroaki Murakami, Reiji Masuda, Tetsuya Sado, Masaki Miya, Sho Sakurai, Hiroki Yamanaka, Toshifumi Minamoto, Michio Kondoh

Data type: Supplementary Table

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/mbmg.2.23297.suppl7

Supplementary material 8

The relationship between MiSeq sequence reads and DNA copy numbers quantified by qPCR

Authors: Masayuki Ushio, Hiroaki Murakami, Reiji Masuda, Tetsuya Sado, Masaki Miya, Sho Sakurai, Hiroki Yamanaka, Toshifumi Minamoto, Michio Kondoh

Data type: Supplementary Figure

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/mbmg.2.23297.suppl8

Supplementary material 9

Bland-Altman plots for the total fish eDNA, Japanese anchovy (*Engraulis japonicus*), and Japanese jack mackerel (*Trachurus japonicus*)

Authors: Masayuki Ushio, Hiroaki Murakami, Reiji Masuda, Tetsuya Sado, Masaki Miya, Sho Sakurai, Hiroki Yamanaka, Toshifumi Minamoto, Michio Kondoh

Data type: Supplementary Figure

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/mbmg.2.23297.suppl9

Supplementary material 10

Abundance of some common fish species obtained by the direct visual census

Authors: Masayuki Ushio, Hiroaki Murakami, Reiji Masuda, Tetsuya Sado, Masaki Miya, Sho Sakurai, Hiroki Yamanaka, Toshifumi Minamoto, Michio Kondoh

Data type: Supplementary Table

Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Link: https://doi.org/10.3897/mbmg.2.23297.suppl10